Post-Hartree Fock Methods - II

2 The method of Configuration Interaction (CI)

Let us now go back to the exact two-electron wave function of equation 10.12 and consider one term (configuration) corresponding to the choice $i=1, j=1$ (the lowest energy Hartree Fock orbital) and neglect all other terms in the expansion. Then

$$\psi(x_1, x_2) = 2c_{11}\{\frac{1}{\sqrt{2}}|\phi_{1\alpha}(1)\phi_{1\beta}(2)|\}$$

$$= a_{11}\Phi_{HF} \tag{13}$$

where $|a_{11}|^2$ give the probability of finding the system in the state described by the ground state HF wave function. All other terms in the expansion in equation 10.12 represent the contributions from the non-HF configurations which naturally arise because of the mutual interaction of the electrons. We note here that all the terms within $\{\}$ in equation 10.11 are singlet eigenfunctions of the S^2, S_z operators and are generally known as spin-adapted configurations (SACs) or configurational spin functions (CSFs). The use of singlet spin adapted configurations in equation 10.11 keeps the number of terms small and avoids need of calculating matrix elements that would have been zero by spin-symmetry.

Note 1. (a) The *n*-electron Hamiltonian (H) is independent of spin in the Schrondinger description and H commutes with S^2 and S_z.

(b) One could use the spatially symmetry adapted CSFs to further reduce the number of non-vanishing matrix elements.

So, using Φ_{HF} as the reference function, we can rewrite equation 10.11 so as to bring home the meaning of different CSFs clearly. Thus we have

$$\psi(x_1, x_2, S=0) = a_{11}\Phi_{HF} + \sum_{i=2} a_{1j} \frac{1}{\sqrt{2}} \{|\phi_{1\alpha}(1)\phi_{j\beta}(2)| + |\phi_{1\beta}(2)\phi_{j\alpha}(1)\}$$

$$+ \sum_{i,j=2} a_{ij} \frac{1}{\sqrt{2}} \{|\phi_{i\alpha}(1)\phi_{j\beta}(2)| + |\phi_{j\alpha}(1)\phi_{i\beta}(2)|\}$$

$$= a_{11}\Phi_{HF} + \sum_{j=2} a_{1j}\phi_{1j}^{(S)} + \sum_{i,j=2} a_{ij}\{\phi_{ij}^{D}\} \qquad (14)$$

$\phi_{ij}^{(S)}$s represent spin singlet configurations which are singly excited with respect to the ground singlet HF configuration (one of the two electrons still resides in the lowest occupied HF orbital (i=1) while the other has been promoted to an orbital ϕ_j with appropriate spin function). Similarly ϕ_{ij}^{D}s represent CSFs built up by promoting two electrons from the occupied HF spin orbitals to appropriate unoccupied spin orbitals (so that singlet spin symmetry is preserved for the CSFs). For a two electron system no other CSFs contribute. In equation 10.14, the amplitude a_{11} is the largest one (≈ 1). $|a_{1j}|, |a_{2j}| \ldots |a_{ij}|s$ are small corrections to the wave function arising out of mixing with available excited state singlet configurations. If doubly excited configurations are dropped, a_{1j}s are all zero as Billouin's theorem asserts that no singly excited configuration function can mix with the variationally determined HF determinant.

Problem 2. Prove that the absolutely best antisymmetrized Product wave function Φ_0 ofan*n*-electron system does not mix with any other *n*-electron antisymmetrized product wave function Φ_1 obtained from the former by replacing one of the orbitals in Φ_0 by an orbital orthogonal to all the orbitals in Φ_0.

In the presence of doubly excited configurations ($S=0$), a_{1j}s acquire small values as the matrix elements of H between the singly and doubly excited configurations are non-zero as also are the matrix elements of H between the HF and the doubly excited determinants. All these amplitudes can be determined by the application of linear variational principle [?].

Problem 3. For a two-electron system construct all the Slater determinants using two space orbitals ϕ_0 and ϕ_1 and the two spin functions α and β. Set up the CI matrix,

starting with the singlet ground state Hartree Fock function.

We are now in a position to generalize the expansion 10.14 for the wave function of *n*-electron system described by a Hartree Fock Roothan wave function built up from LCAO-MO-SCF spin orbitals generated from N_B number of atomic orbital basis functions. N_B basis function will generate N_B spatial HF orbitals of which $\frac{n}{2}$ will be occupied doubly by electrons of opposite spins (each space orbital ϕ_i generates two spin orbitals $\phi_{i\alpha}, \phi_{i\beta}$) ($N_B - \frac{n}{2}$) space orbitals remaining unoccupied at the HF level. The configuration interaction wave function for this general can be written down as

$$\psi_{CI}(S=0) = c_1\Phi_{HF}(S=0) + \sum_S c_S\Phi_S(S=0) + \sum_D c_D\Phi_D(S=0) + \sum_T c_T\Phi_T(S=0) + \ldots$$

$$= \sum_{I=HF,S,D,T,\ldots} c_I\Phi_I(S=0) \qquad (15)$$

The amplitudes C_I s can be determined by minimizing the expectation value of energy $< E >= <\psi_{CI}|H|\psi_{CI}>$ subject to the condition that ψ_{CI} is normalized ($<\psi_{CI}|\psi_{CI}>=1$). With the normalization constraint incorporated through a Lagrange multiplier (λ), we can reformulate the original problem as an equivalent unconstrained one involving minimization of the function f, a Lagrangian so to say, where

$$f = <\psi_{CI}|H|\psi_{CI}> + \lambda[1 - <\psi_{CI}|\psi_{CI}>]$$

$$= \sum_{i=1} c_i^2 E_i + \sum_{i=1}\sum_{j\neq i} c_i c_j <\phi_i|H|\phi_j> + \lambda(1 - \sum_{i=1} c_i^2) \qquad (16)$$

Making f stationary with respect to variations in the unknown CI amplitudes leads to a set of equations equal in number to the number of unknown CI amplitudes (c_i s). Thus setting $\frac{\partial f}{\partial c_i} = 0$ gives us the i^{th} secular equation:

$$\sum_{j\neq 1} c_j <\phi_i|H|\phi_j> + (E_i - \lambda)c_i = 0, \; i = 1,2,3,\ldots,M \qquad (17)$$

Eqn. (17) together with normalization condition $\sum_I c_I^2 = 1$ form a complete set of equations for determining the M unknown energies (E_I) and the corresponding coefficients $c_I(E_I)$. It is convenient to write the secular equations in the form of a matrix eigenvalue equation

$$HC = E\mathbf{1}C \tag{18}$$

where H is an $M \times M$ Hermitian matrix of the Hamiltonian between the CSFs (called the CI matrix) and $E_1 ... E_M$ are the energies of M states and are upper bounds to the corresponding exact eigen energies ($E_i \geq E_i^{exact}, i = 1,2,...M$), $\mathbf{1}$ represents an $M \times M$ identity matrix. How costly is the approach?

For a system with N electron and M basis functions the total number N_c of singlet CSFs that can be generated is given by

$$N_c = \frac{M!(M+1)!}{(\frac{N}{2})!(\frac{N}{2}+1)!(M-\frac{N}{2})!(M-\frac{N}{2}+1)!} \text{(Weyl's dimension formula)} \tag{19}$$

It is easy to see from Eq. 10.21 that N_c grows factorially and even with moderate basis size ($M \approx 20$) and a small number of electrons (10), N_c would be of the order of a billion(N_c grows factorially). The determination of the lowest eigenvalue of the $H-$ matrix (billion × billion) can be done on a modern computer in few hours – but even then the result may still be far from the experimental ground state energy of the molecule. The computational labor required to extract the lowest eigenvalue (or for that matter, one eigenvalue) scales linearly with N_c. So the full or the complete CI calculation is possible only for very small molecules with a small number of electrons and using reasonable basis sets. The full or complete CI eigenvalues, if available, provide the benchmark as these energies are the best possible values that can be obtained with the given basis set. The size of the problem and the exponential growth of computational labour with the size of the CI matrix force one to adopt a truncated CI method in which the expansion for ψ_{CI} of equation 18 is cut-off at S, D or T levels, leading to CIS (single excitation CI), CISD (CI including singly as well as doubly excited configurations), CISDT, etc methods.Once the level of CI has been set the problem reduces to the diagonalization of H matrix which is of the order of million × million. Calculation of all the eigenvalues of such matrices is prohibitively costly. All that one can hope to do is to compute the lowest or a few of the lowest eigenvalues of

the H matrix. The so called direct CI method does this conveniently. In this approach one never computes the full H matrix and stores it. Only a row or a fraction of the terms in a row is calculated and directly multiplied with the corresponding elements of the desired eigenvector C_0 for the ground state. C_0 is unknown to start with; so only a guessed C_0^0 can be provided. Once the full matrix ($N_c \times N_c$) vector ($N_c \times 1$) multiplication has been done the error ($HC_0^0 - E_0 C_0^0$) is computed and based on the error the coefficients are readjusted to generate a new vector C_0^1 following Davidson's method, thereby setting up an iterative process which converges fast to the desired eigenvalue and vector. The matrix-vector multiplication step can be parallelized, leading to a significant reduction in the elapsed-time to get to the desired result.